Supplementary Information – Online Resource 3

Evaluating cost-utility of continuous glucose monitoring in individuals with type 1 diabetes: a systematic review of methods and quality of studies using decision models and/or empirical data.

de Jong LA[1]* (ORCID ID: 0000-0001-8814-0670), Li X[2] (ORCID ID: 0000-0002-0225-6937), Emamipour S[3], van der Werf S[4] (ORCID ID: 0000-0001-5856-7657), Postma MJ[1,5], van Dijk PR[6] (ORCID ID: 0000-0002-9702-6551), Feenstra TL[2] (ORCID ID: 0000-0002-5788-0454)

[1] Department of Health Sciences, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

[2] Unit of PharmacoTherapy, -Epidemiology & -Economics, University of Groningen, Groningen Research Institute of Pharmacy (GRIP), Groningen, the Netherlands

[3] Department of Clinical Pharmacy and Pharmacology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

[4] Central Medical Library, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

[5] Department of Economics, Econometrics and Finance, Faculty of Economics & Business, University of Groningen, Groningen, the Netherlands

[6] Department of Endocrinology. University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

*Corresponding author: t.l.feenstra@rug.nl

Table 1. Interpretation of the Philips et al. 2006 critical appraisal questions as used in the review [1]**.**

Dimension of quality	Questions for critical appraisal	Answered on paper- or model-level?	Interpretation of question for the purpose of this study
S1: Statement of decision problem/objective	1.Is there a clear statement of the decision problem?	Paper level	*Decisions problem is defined here leniently as research/knowledge gap (reason why they're doing this research, so a clear statement of study aim). Please also note the decision problem stated in the study here.*
	2.Is the objective of the evaluation and model specified and consistent with the stated decision problem?	Paper level	*If no decision problem is mention (previous question) then NA. Since 1 was interpreted as requiring an objective of the study, here the interpretation refers mainly to the model specified.*
	3.Is the primary decision maker specified?	Paper level	*Only yes if the authors explicitly mention the decision maker (so this is not the same as the perspective or the study aim). Here we can distinguish the studies that really do consider the actual decision problem supported by the study.*
S2: Statement of scope	4.Is the perspective of the model stated clearly?	Paper level	*State Yes if the perspective was stated (societal, healthcare system, or payer)*
	5.Are the model inputs consistent with the stated perspective?	Paper level	*For healthcare payer perspective: check if all relevant costs related to CGM (and insulin), complications etc were included. For societal costs: often only productivity losses were included, but other societal costs like family members travelling to the hospital (for example in case of severe hypos), and especially informal care costs also matter.*
	6.Has the scope of the model been stated and justified?	Paper level	*This was operationalized as follows: The paper should report what population was included (adults, or young or all ages, everyone or high risk, specific trial population), time-horizon, and the events included (number and type of complications).* *Yes if: if they explain the population, mention time horizon, and complications included.* *The item also asks whether this was justified, which is quite a subjective item to score. Most papers scored stated not justified. We scored this as yes.*
	7.Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?	Paper level	*All will be 'yes', because we included only studies that present an ICER, using QALYs as the outcome.*
S3: Rationale for structure	8.Has the evidence regarding the model structure been described?	Model level	*Yes, if they did a review of relevant complications or existing models, used expert advises, or the model was based on the structure of an existing*

			model. Otherwise, no. (so NR or NA not a possible answer). For CORE model: YES, when they refer to the Palmer et al. and/or McEwan papers.
	9.Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	Model level	*If the model structure excludes a lot of relevant complications (such as hypoglycemia), then no. Otherwise, yes*
	10.Have any competing theories regarding model structure been considered?	Model level	*Yes if any alternatives to the current model structure were described and explained why these were not used.*
	11.Are the sources of data used to develop the structure of the model specified?	Model level	*Yes, if the study reported the sources for the structure (especially if existing model was used). NO if no such information (so not NR in principle). NB large overlap with item 8.*
	12. Are the causal relationships described by the model structure justified appropriately?	Model level	*Yes, if they did validation or if they refer to some studies for this. NB this was a rather lenient interpretation. Very few papers if any were explicit about this. For some more simple models, this could be questioned, especially when it is not possible to get two complications directly from the uncomplicated state, or other strong assumptions regarding combinations of complications.*
S4: Structural assumptions	13. Are the structural assumptions transparent and justified?	Model level	*Yes. If the study reports on the structural assumptions. This implies that we scored yes when they were transparent and did not score whether or not such assumptions were justified. If no assumptions were reported, this was a NO. (not transparent).*
	14. Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Model level	*Matter of judgement, subjective; this was discussed and compared among the reviewers after data extraction was completed. Even after consensus meetings, results varied by duo of reviewers. We added a check to ensure all papers for CORE were scored similarly. As were those based on McQueen/OHTA models.*
S5: Strategies/assumptions	15. Is there a clear definition of the options under evaluation?	Paper level	*If they just say CGM, the answer is no. If it was clear what the answer was in the economic evaluation details tab (including CGM specifics) then yes.*
	16. Have all feasible and practical options been evaluated?	Paper level	*Yes, if based on a review of the existing evidence at the time of the study. Y/N when clearly based on the extrapolation of a single trial so practically the trial determines the options. NO else.*
	17. Is there justification for the exclusion of feasible options?	Paper level	*Yes, if they discuss other feasible options (see previous question) and give an explanation why these were not included. No if other options were not discussed/justified. (cannot be NR here). IF 16=YES, this has to be NA.*
S6: Model type	18. Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?	Model level	*Yes if they used a State transition model (patient level or cohort level). Hard to judge at the more detailed level.*

S7: Time horizon	19. Is the time horizon of the model sufficient to reflect all important differences between options?	Paper level	*Only yes if long-term time horizon (such as lifelong) was used. Else NO*
	20. Is the time horizon of the model, and the duration of treatment and treatment effect described and justified?	Paper level	*Overlaps a lot with other items. Here we want to answer the question if the CGM effect duration is transparent and justified. Answered YES when clear what was done regarding duration of treatment effects. NO else. (Many did not report, that is a NO)*
	21. Has a lifetime horizon been used? If not, has a shorter time horizon been justified?	Paper level	*Overlaps with previous two questions, make sure answers are aligned. Yes if lifetime horizon was used, no if not.*
S8: Disease states/pathways	22. Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	Model level	*It's only 'no' if the conceptual model is considered not representative for the disease. Hard to judge. Some models are super simple and preclude having multiple comorbidities by structure. That should be a no. Overlap with previous items on model structure.*
S9: Cycle length	23. Is the cycle length defined and justified in terms of the natural history of disease?	Paper level	*Yes if cycle length was mentioned. One year cycles were considered appropriate to model diabetes-related complications. Hypoglycemic events should be modelled with a shorter cycle length (for example three months) or in background to make sure these can occur more than once per year.*
D1: Data identification	24. Are the data identification methods transparent and appropriate given the objectives of the model?	Paper level	*Only yes if a review was conducted to identify input data and were appropriately reported in the specific paper. This was only rarely done. If no information, then a NO, since not transparent.*
	25. Where choices have been made between data sources, are these justified appropriately?	Paper level	*Yes if the specific paper provides justification for choice being made between input data. Most papers just use single sources, without further clarification. In that case would be NR (no information).*
	26. Has particular attention been paid to identifying data for the important parameters in the model?	Paper level	*NR if they don't mention anything about extra efforts they have done to identify the most important parameters. If yes, please provide methods used and which parameters where considered important.*
	27. Has the process of selecting key parameters been justified and systematic methods used to identify the most appropriate data?	Paper level	*Only yes if a review was conducted to identify data for important input parameters, like effectiveness. Seems large overlap in type of item with 26. NO when not done.*
	28. Has the quality of the data been assessed appropriately?	Paper level	*Yes if quality assessment was conducted. This is very hard to check from the papers. Also not discussed in Palmer or McEwan. NR else.*
	29. Where expert opinion has been used, are the methods described and justified?	Paper level	*Only external expert opinion is relevant to this question, not expertise by the authors themselves. This concerns expert opinion to fill data gaps. If no information on how it is a NO (since asks for description). If no expert opinion, it can be NA.*

D2: Pre-model data analysis	30. Are the pre-model data analysis methodology based on justifiable statistical and epidemiological techniques?	Paper level	*NA, if they don't do statistics. Yes, if trial/meta-analysis was conducted specifically for this model and if they used justifiable statistical and epidemiological techniques. Very often, reference is made to a background model report, which will contain some (but not all information) on pre-model data analysis. In this case score NR, not NA, since data-analysis was applied.*
D2a: baseline data	31. Is the choice of baseline data described and justified?	Paper level	*Yes if they provide the source, inputs and references. No specific justification for the source needed. This refers to baseline population data. (so characteristics of population). NO if not clear or no description.*
	32. Are transition probabilities calculated appropriately?	Paper level	*NR if they do not provide calculations for transition probabilities, prediction models, or risk equations. Or if they do not give a reference to a background paper that has this.*
	33. Has a half cycle correction been applied to both cost and outcome? If not, has this omission been justified?	Paper level	*Yes if half cycle correction has been applied or if omission is justified.*
D2b: treatment effects	34. If relative treatment effects have been derived from trial data, have they been synthesized using appropriate techniques?	Paper level	*NA if they do not synthesize. Yes, if they use systematic review or meta-analysis to identify and select relative treatment effect. NA, when a single RCT is used.*
	35. Have the methods and assumptions used to extrapolate short-term results to final outcomes been documented and justified? Have alternative assumptions been explored through sensitivity analysis?	Paper level	*Yes if methods for extrapolation of short term results to final outcomes have been reported. Not so sure what is meant by this. In principle this is the complete model-based evaluation (extrapolate from short term to QALYs and costs). Scored yes when SA and description of how treatment effects went into the model was clear enough. NO when this was very unclear and/or no SA at all.*
	36. Have assumptions regarding the continuing effect of treatment ~~once treatment is complete~~ been documented and justified? Have alternative assumptions been explored through sensitivity analysis?	Paper level	*CGM is in principle a lifelong treatment. Yes if it was clear whether continuing effects were assumed. NO if this was unclear. Almost no SA on this, so scored already yes when clear discussion/clarity on assumptions.*
D2c: quality-of-life weights (utilities)	37. Are the utilities incorporated into the model appropriate?	Paper level	*For hypoglycemia: based on assumptions is not considered appropriate. FoH score should be based on a study. And explained what was done. Diabetes-related complications: at least a reference should be provided. Else NR.*
	38. Is the source for the utility weights referenced?	Paper level	*Yes, if the source for each specific utility weight (for each health state/complication) was reported. Else NO.*

	39. Are the methods of derivation for the utility weights justified?	Paper level	*No if the specific method was not mentioned in the paper. Yes if they conducted a systematic review and make a clear choice which utility they use. Or if they conduct quality of life study themselves and methods are appropriate. Yes when any mapping (for FOH often) was clearly explained and was not just based on some assumption.*
D3: Data incorporation	40. Have all data incorporated into the model been described and referenced in sufficient detail?	Paper level	*Yes if they provide sources in text or in tables (or supplementary materials). No if unclear which sources were used for which parameters.*
	41. Has the use of mutually inconsistent data been justified (i.e. are assumptions and choices appropriate)?	Paper level	*NR if this is not discussed in the paper. Yes if they make an explicit choice between mutually inconsistent data and do a sensitivity analysis to explore the effect of the input data choice.*
	42. Is the process of data incorporation transparent?	Paper level	*No if they provide just a source for an input parameter without giving any justification. Choice must be made clear. (NO or YES)*
	43. If data have been incorporated as distributions, has the choice of distribution for each parameter been described and justified?	Paper level	*If they describe the distribution for each input type, justification is not needed. NA when no PA.*
	44. If data have been incorporated as distributions, is it clear that second order uncertainty is reflected?	Paper level	*Yes if it was clearly stated a second-order Monte Carlo simulation or analysis of parameter uncertainty was conducted. Else NA (if no CEAC) or NO (When CEAC, but not very clear how)*
D4: Assessment of uncertainty	45. Have the four principal types of uncertainty been addressed?	Paper level	*Methodological* *Structural* *Heterogeneity* *Parameter* *Only 'yes' if all four have been assessed, otherwise no. So should be consistent with 43/44 and with 47-51. .*
	46. If not, has the omission of particular forms of uncertainty been justified?	Paper level	*If previous answer was no, please indicate if omission was justified. If previous answer was yes, answer this question with NA.*
D4a: methodological	47. Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	Paper level	*Yes if for example scenario have been run on time horizon or discount rates.*
D4b: structural	48. Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	Paper level	*Yes if any sensitivity analyses have been conducted to assess the influence of an alternative model structure or adding or leaving out certain parts of the model structure used in the base case analysis.*
D4c: heterogeneity	49. Has heterogeneity been dealt with by running the model separately for different sub-groups?	Paper level	*Yes if they performed subgroup analysis or any sensitivity analysis using alternative patient characteristics.*
D4d: parameter	50. Are the methods of assessment of parameter uncertainty appropriate?	Paper level	*Yes if the methods for obtaining the ranges/confidence intervals for the input parameters was appropriate. Either PA or SA.*

	51. Has probabilistic sensitivity analysis been done, if not has this been justified?	Paper level	*Yes if PA has been conducted. No if this has not been conducted and please report if the paper justified omission. Almost never justified. NO (not NR) when no PA. Should be consistent with 43 and 44.*
	52. If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?	Paper level	*Yes if they describe ranges for univariate sensitivity analysis. NO if not.*
C1: Internal consistency	53. Is there evidence that the mathematical logic of the model has been tested thoroughly before use?	Model level	*Yes if mathematical logical was tested/validated and this was reported in the specific paper or the background paper(s) of the model used.*
C2: External consistency	54. Are the conclusions valid given the data presented?	Paper level	*Yes if the conclusion about cost-effectiveness results was correct given the ICER at the WTP-threshold and based on the sensitivity analysis.*
	55. Are any counterintuitive results from the model explained and justified?	Paper level	*Yes, if the specific paper discusses any counterintuitive results in the discussion section. NA if there were no counterintuitive results.*
	56. If the model has been ~~calibrated~~ validated against independent data, have any differences been explained and justified?	Model level	*Yes if this has been reported in the specific paper or in one of the background paper(s) they refer to for the model used. The answer should be in line with the answers from the AdViSHE checklist (item 12).*
	57. Have the results of the model been compared with those of previous models and any differences in results explained?	Model level	*Yes if this has been reported in the specific paper or in one of the background paper(s) they refer to for the model used. The answer should be in line with the answers from the AdViSHE checklist (item 10). Should be Yes or NO, NR or NA is not possible.*

Abbreviations: CEAC, cost-effectiveness acceptability curve; CGM, continuous glucose monitoring; FoH, fear of hypoglycemia; ICER, incremental cost-effectiveness ratio; NA, not applicable; NR, not reported; PA, probabilistic analysis; WTP, willingness to pay.

References

1. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics. 2006;24(4):355–71.